MANUSCRIPT
WRE # 366

Statistical Modeling of Dry Deposition Phosphorus Rates Measured from South Florida

Submitted for publication in Environmental Pollution

August 1998

By

Hosung Ahn
Water Resources Evaluation Department
R. Thomas James
Ecosystem Restoration Department

SOUTH FLORIDA WATER MANAGEMENT DISTRICT
3301 Gun Club Road
West Palm Beach, Florida 33406

Title: **Statistical Modeling of Dry Deposition Phosphorus Rates Measured from South Florida**

Authors: Hosung Ahn and R. Thomas James

Date: August 25, 1998

Target: Environmental Pollution

Summary:

Dry atmospheric deposition may be a substantial source of phosphorus (P) to the Everglades. Measured on a weekly basis at nineteen sites throughout the District, samples often are discarded because of contamination from bird droppings and other foreign material. This can create large gaps in the time line of the data sets. Missing data were so extensive at six stations, either in the amount or in the size of gaps, that they were excluded from further analysis. For the remaining thirteen stations, we estimated the missing data with statistical models. These models calculate values of missing samples at a given site based on relationships to previous samples at that given site and to current samples at nearby sites.

The estimated data are quite accurate. The overall mean and standard deviation of the data before estimating the missing values was 88.4 ± 85.7 µg P m^{-2} d^{-1} and after estimating the missing values it was 87.8 ± 82.4 µg P m^{-2} d^{-1}. For each sampling site the mean and standard deviation before and after were quite similar. No trend with time was detected. The P values fluctuate seasonally (highest in October and lowest in June) but this fluctuation does not follow the seasonal pattern of South Florida's rainfall. Random noise in the data, however, was significant and caused long-term fluctuations of the data. The data after estimating missing values are useful for accurately calculating the weekly P loads from atmospheric deposition.

Statistical Modeling of Dry Deposition Phosphorus Rates

Measured from South Florida

Hosung Ahn[1] & R. Thomas James[2]

[1]Water Resources Evaluation Department

and

[2]Ecosystem Restoration Department

South Florida Water Management District

3301 Gun Club Road

West Palm Beach, Florida 33406

Tel: (561) 687-6516

Fax: (561) 687-6442

Email: hosung.ahn@sfwmd.gov

Abstract

Dry atmospheric deposition may be a substantial source of phosphorus (P) to the Florida Everglades. Dry deposition has been measured on a weekly basis in the region since 1987, but a significant amount of this data is missing (about 34 percent) due to instrumental failures and sample contamination. This study develops a set of statistical models of the P dry deposition time-series to estimate missing data. These models are based on a multivariate stochastic time-series theory. Model parameters and noise covariances are calibrated using the expectation-maximization algorithm which is efficient for data sets with many gaps. The pooled mean and standard deviation of the data before estimating the missing values was 88.4±85.7 μg P m^{-2} d^{-1} and after estimating the missing values was 87.8±82.4 μg P m^{-2} d^{-1}. Model verification demonstrates that the calibrated models provide unbiased data estimates while preserving the statistics of the raw data. For each sampling site the mean and standard deviation before and after were quite similar. No trend with time was detected. The P deposition fluctuates seasonally (highest in October and lowest in June), but this fluctuation does not follow the seasonal pattern of South Florida's rainfall. Random noise in the data, however, is significant and causes long-term fluctuations of the data. The data with gaps filled in are useful for computing the weekly P load distribution.

Key words: Dry deposition; phosphorus load; missing data; time-series model; Kalman filtering and smoothing.

Introduction

Increased phosphorus concentrations have resulted in significant changes to the Everglades ecosystem (Davis, 1994). As a result the State of Florida enacted the Everglades Forever Act (State of Florida, 1994) that mandates the management of phosphorus loads to the south Florida ecosystem through the use of stormwater treatment areas and best management practices. Non-controllable sources such as atmospheric deposition also must be accurately monitored to assess their impact in relation to all other loads.

Atmospheric deposition is measured in wet (rainfall) and dry (dustfall) forms. Samples of the latter are often contaminated and have to be removed from any analysis. Gaps created from this removal of data preclude accurate calculation of weekly P loads. Monthly or yearly summary statistics can be affected by gaps that increase the influence of non-missing samples in the calculation. This is important because dry deposition is quite variable over time (Hicks et al., 1993). Given a single high or low sample in a given month surrounded by missing data, the average for that month may be skewed higher or lower because of that one high or low sample. To achieve an accurate account of this variability, missing samples should be estimated.

If the physical processes driving the occurrence and transportation mechanism of atmospheric deposition are known, one could build a mathematical model to estimate the data gaps. However, neither a mathematical model nor supporting input data for the model on a regional scale is available for south Florida. Alternatively, we selected an empirical approach using simple statistical models based on currently available P deposition rates. Such an approach has been used for wet deposition (Ahn 1998), but to our knowledge, it has not been used for dry deposition. Our objective was to develop statistical models to estimate the missing P values. This particular procedure could be used in any multi-site atmospheric deposition study.

Statistical Methods to Estimate Missing Data

This study uses a multi-site time-series model. Numerical algorithms such as Gauss-Newton or scoring method (Box and Jenkins, 1976; Brockwell and Davis, 1987; Harvey, 1990) can estimate parameters in the time-series models, but they are not applicable for incomplete data sets. An expectation-maximization (EM) algorithm is suitable for estimating parameters of time-series models for data sets with gaps (Dempster et al., 1977; Shumway and Stoffer, 1982; Stoffer, 1985; Stoffer, 1986). A pre-condition to applying the EM algorithm is to set the model into state-space form to estimate Kalman filtering and smoothing estimates. The Kalman filter and smoothing methods provide a convenient means to calculate the conditional expectations of both state and error vectors. The reason that smoothing is used is to take advantage of the forward measurement information and to give a fast convergence in the EM algorithm.

Consider a multi-site state vector $x_t=(x_{t,1},...,x_{t,nx})'$ at time t (=1,...,T), where (nx) is the number of sites of the state variable, T is the time span, and ' denotes transpose of a matrix. With the (nz) multi-site covariate vector $z_t=(z_{t,1},...,z_{t,nz})'$ which is measured completely and

$$x_t = \sum_{i=1}^{q} \phi_i x_{t-i} + \psi z_t + w_t \tag{1}$$

concurrently, the order-q multi-site autoregression model is given by,

where $\phi_i(nx \times nx)$ and $\psi(nx \times nz)$ are the regression parameters, and $w_t(nx \times 1)$ is the Gaussian white noise with $w_t \approx N(0,Q)$. For P dry deposition data, x_t could represent a P vector measured from (nx) multiple sites at time t, while z_t may be a concurrently measured covariate vector having a size of (nz).

To estimate the parameters of the above model with incomplete data, an EM algorithm is applied in conjunction with the modified Kalman smoother estimators. To apply the Kalman filter, Eq. 1 should be set into a state-space form that consists of state and measurement equations. Using Eq. 1 as a state equation, the measurement equation that allows for missing data can be written as,

$$y_t = m_t x_t + v_t \tag{2}$$

where y_t $(nx \times 1)$ is the state measurement vector at t, m_t is the $(nx \times nx)$ measurement matrix in which the diagonal element is 1 if y_t is measured, or 0 otherwise, and v_t is the Gaussian white noise having $v_t \approx N(0,R)$. In this state-space problem, the parameter set to be calibrated is $\{\Phi, \psi, Q, R\}$ with $\Phi = [\phi_1, \ldots, \phi_q]$. An incomplete-data log-likelihood function, $2\ln L(y)$, is calculated (Gupta and Mehra, 1970; and Shumway and Stoffer, 1981) by,

$$2\ln L(y) = \sum_{t=1}^{T} \ln|m_t\, p_t^{t-1} m_t'| + \sum_{t=1}^{T} e_t^{t-1\prime} [m_t\, p_t^{t-1} m_t' + R]^{-1} e_t^{t-1} \tag{3}$$

where $e_t^{t-1} = (y_t - m_t x_t^{t-1})$ is the measurement error, and $x_t^{t-1} = E[x_t | x_1, \ldots, x_{t-1}, z_1, \ldots, z_t]$ is the state estimate conditioned on prior information and p_t^{t-1} is the corresponding error covariance term

(Jazwinski, 1970).

The iterative procedure of the EM algorithm (Shumway and Stoffer, 1981; Shumway, 1988) starts with an assumed initial parameter set, $\{\Phi(0), \psi(0), Q(0), R(0)\}$, where (0) indicates the initial time step before iterations. On the i-th iteration, the Kalman filtered and smoothed estimators with the log-likelihood function (Eq. 3) are computed. The procedure continues until the log-likelihood function converges.

Data Collection and Analysis

The South Florida Water Management District (District) has collected atmospheric deposition data in this region since 1974. Before 1987 only bulk collectors were used. The monitoring program was significantly improved in 1987 with the deployment of wet/dry collectors (Aerochem Metrics Model 301 automatic wet/dry sampler) and adoption of a standard operating procedure for atmospheric data collection and processing. Currently, there are 19 atmospheric deposition monitoring sites operated by the District (Figure 1). Wet and dry deposition data have been collected at weekly intervals and analyzed at the District's laboratory to determine the level of nutrients and major ions. A complete description of the collection and handling of the samples is found in Ahn and James (1998) and in the District's Comprehensive Quality Assurance Plan (SFWMD, 1996).

Briefly, the dry bucket is inspected for contamination, and any observed contaminants are noted and removed with tweezers, if possible. The bucket is rinsed with 1 L of deionized water and the sides of the bucket are rubbed with a precleaned spatula. The sample is placed into 150 ml Nalgene bottles and acidified to a pH of 2.0 with a 50 percent solution of reagent grade H_2SO_4. The bottles are placed on ice and returned within a day to the District's laboratory where the samples are analyzed for total P with a persulfate digestion method (USEPA 1979).

Results and Discussion

Identification of Model Structure

Among 19 atmospheric deposition monitoring sites, six sites (ENR101, ENR203, ENR301, ENR401, S127, S308) were excluded because of excessive data contamination that caused severe data gaps and divergence in Kalman filtering and smoothing: The average missing percentage of dry P data at these six sites was about 42%. This study used the weekly dry deposition P loads collected from 13 other monitoring sites for the maximum period of record from April 1992 to November 1996. The actual periods of record vary from site to site due to periodic expansion of the monitoring program.

Generally, it would be easy to build one lumped multivariate time-series model for all 13 sites, making it more efficient to estimate model parameters and values for missing data. However, developing such a lumped model in this case was not possible because of the varying periods of record and the lack of covariate information other than the P data concurrently measured from adjacent sites. Because of these limitations, this study constructed five separate models (Table 1).

In the time-series model described by Eq. 1, the state vector x_t at time t is a function of both the time-lagged state x_{t-q} and the concurrently measured covariate z_t. The concurrently measured dry P data collected from sites adjacent to that being modeled were used as the covariates in each model. Because inter-site correlation of concurrently measured data is stronger than auto-correlation (correlation in time), selecting sites for the z_t vector is very important.

Because gaps are not allowed in the z_t vector, the model structure was designed to estimate model parameters and missing data in x_t sequentially by taking the state estimates (with filled-in data) of the previous model (Table 1) and applying it to z_t of the current model. Considering the cross-covariance, periods of record, and the distance from the state site, several alternative models with different combinations of covariates were tried from which an optimal model was selected for each case by maximizing the log-likelihood function of Eq. (3). In order to obtain a complete covariate data set for Model I, Model II without the z_t term was initially used to estimate the missing data in $x_t' = [x_{okeefs,t}, x_{S140,t}]$.

The order q in Eq. 1 was determined using Akaike Information Criteria (AIC) (Shumway, 1988) which chooses the model order q that minimizes,

$$AIC(q) = \ln\left(\sum_{t=1}^{T}[w_t' w_t]/T\right) + 2nx^2 q/T \qquad \textbf{(4)}$$

For each model, alternative models having different orders ranging from q=1 to q=4 were tried for identification of the best model order. Based on the AIC statistics, q=1 was dominant in all five models. During the model identification process, the low-order model (q=1) was preferred because increasing the number of model parameters adversely affected the AIC statistics.

Parameter Calibration

The large number of missing samples in the measured data sets were caused by sample contamination due to bird droppings and other foreign matter: An average of 34 (ranging from 10 to 58) percent of dry deposition P data collected on a weekly basis in the region was missing from the historical data sets (Table 2).

The distribution of raw data were positively skewed (with skewness coefficients ranging from 0.23 to 1.05), and were log-transformed before the models were applied. After setting up the measurement matrix in each model based on the availability of data, the parameters of the five models were sequentially estimated using the EM algorithm. For example, Model II where $x_t'=[x_{t,okeefs}, x_{t,S140}]=[x_{1,t}, x_{2,t}]$ and $z_t'=[z_{ENR,t}, z_{S65A,t}, z_{S7,t}]=[z_{1,t}, z_{2,t}, z_{3,t}]$ are the log-transformed air-borne P rates in $\mu g\ m^{-2}\ d^{-1}$, and T=240, the calibrated model is given by

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} 0.722 & -0.328 \\ -0.321 & 0.022 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} 0.900 & 0.422 & 0.356 \\ 0.354 & 1.124 & 0.669 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{3,t} \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix} \tag{5}$$

with the diagonal terms in Q, which represents noise (Eq. 1), and R, which represents measurement equations (Eq. 2), matrices being [1.936, 1.599] and [0.0046, 0.0056], respectively. The regression coefficients for z_t are higher than those of x_t. That is, the inter-site correlation of concurrently measured P values are higher than the time-lagged correlation of the data.

The parameter calibration demonstrated that the smaller the size of missing data, the faster the EM algorithm converges. As a result, the model for a small state dimension (probably 2-4 sites) gives a more accurate estimate for missing data than a larger one. This fact also justified development of five separate models instead of one lumped model. The initial assumed parameter set $\{\phi(0),\psi(0),Q(0),R(0)\}$ needed for the EM algorithm was not sensitive to the final estimation result. This insensitivity was another advantage of the EM algorithm as a parameter estimation method for a time-series model.

Summary Statistics and Trends

The final data sets consisted of direct observations if they were available and model estimates if they were missing (Figure 2). The data varied from near 0 to over 450 μg P m^{-2} d^{-1}. The estimated missing data also reflect this variability. A comparison of the mean and standard deviation before and after filling in the gaps demonstrates that these summary statistics were well preserved (less than 5 percent error) after gaps in the data were filled, with the expected means and standard deviations of the data falling slightly below the 1:1 line with the observed statistics (Figure 3).

This variability in dry deposition has been documented for both space (Hicks et al. 1993, Van Ek and Draaijers 1994, Dixon et al. 1996, Hendry et al. 1981), and time (Hicks et al. 1993). The latter is primarily a result of episodic events and deposition of larger (> 2 μ) particles. Both the spatial and temporal variability are also demonstrated in the District's network of atmospheric deposition stations (Figure 2, Table 2). The standard deviation for each site is equivalent to the mean. Also the means ranged from an average of 14.4 μg P m^{-2} d^{-1} at a remote station in a marsh area of the Everglades (L67A) to 154.2 μg P m^{-2} d^{-1} at S65A a site surrounded by improved pasture. The pooled mean for the 13 sites before filling-in the data gaps was 88.4±85.7 μg P m^{-2} d^{-1} and after filling-in data gaps was 87.8±82.4 μg P m^{-2} d^{-1} (Table 2).

The average estimation errors in the last two columns of Table 2 were computed from the square root of the smoothed error covariance computed from the EM algorithm. The average estimation error for the missing part was about 50 percent greater than that of the overall data (both missing and non-missing parts) but the estimation errors were relatively small compared to their mean (about 6%) implying this estimation of missing data was quite accurate.

Estimates of atmospheric P deposition from dry deposition range from 4 to 10 times that of wet deposition (Hicks et al. 1993). Wet deposition in south Florida, with a mean rainfall of 1.35 m $year^{-1}$ and an mean concentration of 10.6 μg P L^{-1} (Ahn 1998) in rainfall, is estimated as 14.3 mg P m^{-2} $year^{-1}$. Our pooled estimate of dry deposition with estimates of missing data is 32.1 mg P m^{-2} $year^{-1}$. Thus, the ratio of our dry deposition to wet deposition is approximately 2:1, which is lower than others have observed.

The total estimate of wet deposition plus dry deposition, 46.4 mg P m^{-2} $year^{-1}$, is consistent with estimates from peat accretion data of 35.5 mg P m^{-2} $year^{-1}$ (Walker 1993), and 50 mg P m^{-2} $year^{-1}$ from bulk collectors throughout Florida (Hendry et al. 1981). But it is less than the 93.3 mg P m^{-2} $year^{-1}$ determined in the Tampa area from seven bulk collectors (Dixon et al., 1996). These comparisons provides a certain level of confidence regarding the District's sampling network, procedures, and the statistical approach that we have taken.

To investigate the seasonality in the data, the monthly dry P deposition rates from all 13 sites were pooled, and the mean and standard deviation for each month of the year were computed (Table 3a). The monthly mean P values are lowest in June (about 59.7 μg P m^{-2} d^{-1}) and highest in October (120.8 μg P m^{-2} d^{-1}) before filling in the data gaps. This same seasonal trend was observed after filling in the data gaps but the values were slightly lower (58.6 and 110.8 μg P m^{-2} d^{-1} respectively) (Table 3b). The standard deviations were also smaller as were the minimum values. However, after estimating missing data, the maximum values in some months increased. There is a month-to-month fluctuation in the monthly P values, but the fluctuation does not follow the wet-dry rainfall pattern in South Florida where the wet season extends from mid-May to October.

The overall 12 month mean (80.3$\pm$ 14.8 μg P m^{-2} d^{-1}, Table 3a) was smaller than the overall site mean (82.2$\pm$33.6 μg P m^{-2} d^{-1}, Table 2). The difference between these means is a reflection of the differing amounts of data collected at each station. The differences between the standard deviations of the mean for these two values is indicative of a greater spatial than seasonal variation in the data.

Plots of the monthly average time-series of the P deposition of three arbitrarily selected sites shows that there is no temporal trend in the data during the period of record as evidenced by the slopes of the regression lines that are not significantly different from zero (Figure 4). A 6-month moving average used to indicate seasonal trends simply fluctuates due to abnormal high P rates that appear randomly in time. The other sites have the same temporal patterns but are not presented here.

Summary

Samples of dry atmospheric deposition to South Florida are often contaminated and must be discarded. This produces numerous gaps in data sets that measure dry P deposition. We estimated values for this missing information with statistical models. Five multivariate time-series models were developed from historical data collected from 13 monitoring sites. The model parameters and the missing data were estimated simultaneously by an expectation-maximization algorithm. In order to compute the expectation step, the time-series model was set into a state-space form and the Kalman filtering and smoothing algorithms were applied.

As a verification of the model estimates, the statistics of the data were computed after estimating the missing data and compared with those from the original data set. The results were quite satisfactory in that the mean and standard deviation of the data (after estimating the missing data) were preserved. The averages of the estimated means from the 13 sites were 82.2±70.5 μg P m^{-2} d^{-1}, with an average estimation error of 3.1 μg m^{-2} d^{-1}. There were no temporal trends. The P values are high in October and low in June, but the fluctuation did not follow the wet-dry rainfall pattern in South Florida. Instead, random noise in the data appeared to be the main cause of long-term irregular fluctuations in the data. In general, the inter-site correlation of the data was stronger than temporal correlation.

Undoubtedly, the P rates resulting from applying this methodology to estimate missing data can be useful for calculating the weekly P load input from the atmosphere. Alternatively, the load could be calculated for a longer time interval (monthly or yearly), but it would be less accurate than weekly since the temporal variability of the data is very significant.

Acknowledgements.

The authors are grateful to Cheol Mo and Maria Loucraft-Manzano for discussions early in the work, and to Todd Tisdale for constructive comments on the manuscript.

References

Ahn, H., 1998. Estimating the mean and variance of censored phosphorus concentrations in Florida rainfall. Journal of the American Water Resources Association 34: 583-594.

Ahn, H., and R. T. James, 1998. Outlier detection in dry deposition phosphorus rates measured from South Florida rainfall (in preparation).

Box, G. E. P. and Jenkins, G. M., 1976. Time Series Analysis: Forecasting and Control. Prentice-Hall, Englewood Cliffs, New Jersey.

Brockwell, P. J. and Davis, R. A., 1987. Time Series Theory and Methods. Springer-Verlag, New York.

Davis, S. M. 1994. Phosphorus inputs and vegetation sensitivity in the Everglades. In: Everglades--the Ecosystem and Its Restoration. by S. M. Davis and J. C. Ogden (Editors), pp. 357-378, St. Lucie Press, St. Lucie, FL.

Dempster, A. P., Laird, N. M. and Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. B 39: 1-38.

Dixon, L.K., S. Murray, J. S. Perry, P. J. Minotti, M. S. Henry, and R. H. Pierce, 1996, Assessment of bulk atmospheric deposition to the Tampa Bay Watershed, Final Report submitted to the Tampa Bay National Estuary Program, St. Petersburg, FL.

Gupta, N. K., and Mehra, R. K., 1970. Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculation. IEEE Trans. on Automatic Contl., AC-19: 774-783.

Harvey, A. C., 1990. The Econometric Analysis of Time Series. Second Edition, The MIT Press, Cambridge, Massachusetts.

Hendry C.D., Brezonik P. L., and Edgerton, E. S., 1981. Atmospheric deposition of nitrogen and phosphors in Florida. In: Atmospheric Pollutants in Natural Waters, by S. J. Eisenreich (Editor). pp. 199-215, Ann Arbor Sci., Publ., Ann Arbor, MI.

Hicks, B., McMillan, R., Turner, R. S., Holdren, Jr., G. R., an Strickland, T. C., 1993. A national critical loads framework for atmospheric deposition effects assessment: III. Deposition Characterization. Environmental Management, 17: 343-353.

Jazwinski, A. H., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.

SFWMD, 1996. SFWMD comprehensive quality assurance plan. WRE-346, South Florida Water Management District, West Palm Beach, FL.

Shumway, R. H., 1988. Applied Statistical Time Series Analysis. Prentice Hall, Englewood Cliffs, New Jersey.

Shumway, R. H. and Stoffer, D. S., 1981. Time series smoothing and forecasting using the EM algorithm. Tech. Report No. 27, Div of Statistics, University of California, Davis, California.

Shumway, R. H. and Stoffer, D. S., 1982. An application of time series smoothing and forecasting using the EM algorithm. J. Time Series Anal. 3: 253-264.

State of Florida, 1994. Everglades Forever Act.

Stoffer, D. S., 1985. Maximum likelihood fitting of STARMAX models to incomplete space-time series data. In: Time Series Analysis: Theory and Practice, by O. D. Anderson, J. K. Ord, and E. A. Robinson (editor), North-Holland, Amsterdam: 283-296.

Stoffer, D. S., 1986. Estimation and identification of space-time ARMAX models in the presence of missing data. J. American Statistical Assoc. 81: 762-772.

USEPA, 1979. Methods for Chemical Analysis of Water and Wastes, United States Environmental Protection Agency, Washington, D. C.

Van Ek, R. and Draaijers, G. P. J., 1994. Estimates of atmospheric deposition and canopy exchange from three common tree species in the Netherlands, Water, Air, and Soil Pollution 73, 61-82.

Walker, W. W. 1993. A mass-balance model for estimating phosphorus settling rate in Everglades Water Conservation Area 2A. Prepared for the U. S. Department of Justice. 20 pp.

Table 1. Composition of time-series models and periods of record of the historical data used for model calibration.

Model Number	Sites	Periods of Record (month/day/year)	N	Covariate Sites
I	ENR, S65A, S7	4/7/92-11/5/96	240	OKEEFS, S140
II	OKEEFS, S140	4/7/92-11/5/96	240	ENR, S65A, S7
III	BG1, ENPRC, S131	9/7/92-11/5/96	166	ENR, S65A, S7
IV	BG2, S310, G36	9/7/93-11/5/96	166	S65A, ENPRC, BG1
V	L67A, L6	11/21/95-11/5/96	51	ENR, ENPRC

Table 2. Summary statistics of dry deposition of P ($\mu g\ m^{-2}\ d^{-1}$) measured at 13 sites in south Florida. *Ratio of missing points to available data points during period of record. + Standard deviation of column values.

Site	Missing Ratio*	Mean	S.D.	Estimation Error	
				Overall	Missing
BG1	0.26	83.5	80.8	2.9	5.1
BG2	0.17	82.0	85.6	2.6	4.7
ENPRC	0.58	66.6	75.3	4.5	6.1
ENR	0.10	88.3	80.3	2.5	5.1
G36	0.50	134.0	81.6	3.2	4.1
L67A	0.43	14.4	10.2	3.0	3.9
L6	0.35	64.1	49.2	3.0	4.5
OKEEFS	0.44	72.0	70.3	3.6	5.3
S131	0.31	80.7	74.4	3.2	5.3
S140	0.40	67.6	63.6	3.4	5.3
S310	0.31	89.0	68.0	2.8	4.0
S65A	0.42	154.2	96.8	3.3	4.7
S7	0.15	72.4	79.8	2.6	5.1
Mean	0.34	82.2± 33.6+	70.5± 21.4+	3.1	4.9
Pooled Mean with estimates		87.8	82.4		
Pooled Mean without estimates		88.4	85.7		

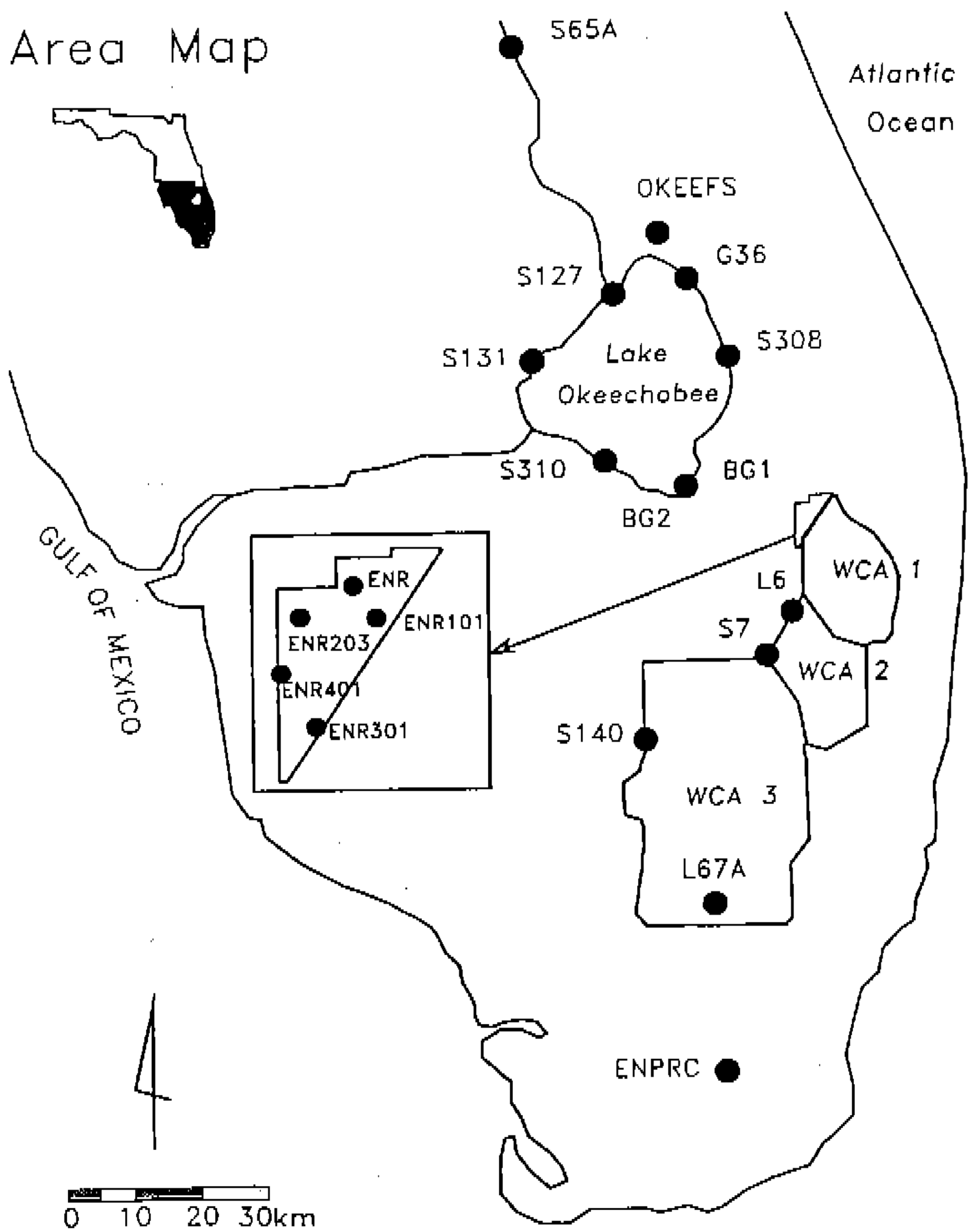
Table 3a. Summary statistics of the thirteen station dry deposition network in South Florida, Pooled by month before filling in data gaps. + Standard deviation of column values. *Pooled maximum or minimum.

Month	Mean	Standard Deviation	N	Minimum	Maximum
January	77.0	63.1	118	4.1	294.1
February	99.3	82.0	104	8.1	363.0
March	95.9	86.2	103	12.2	354.9
April	72.6	66.4	110	8.1	314.3
May	90.2	87.5	116	4.1	361.0
June	59.7	72.3	114	4.1	352.9
July	72.0	73.6	108	4.1	363.0
August	73.2	81.2	131	4.1	354.9
September	89.7	88.0	126	4.1	346.8
October	120.8	105.4	152	4.1	365.0
November	102.8	88.9	110	4.1	354.9
December	101.2	99.4	102	4.1	365.0
Mean	87.9± 17.3+	82.4± 12.6+	116	4.1*	365.0*

Table 3b. Summary statistics of the thirteen station dry deposition network in South Florida, Pooled by month after filling in data gaps. + Standard deviation of column values. *Pooled maximum or minimum.

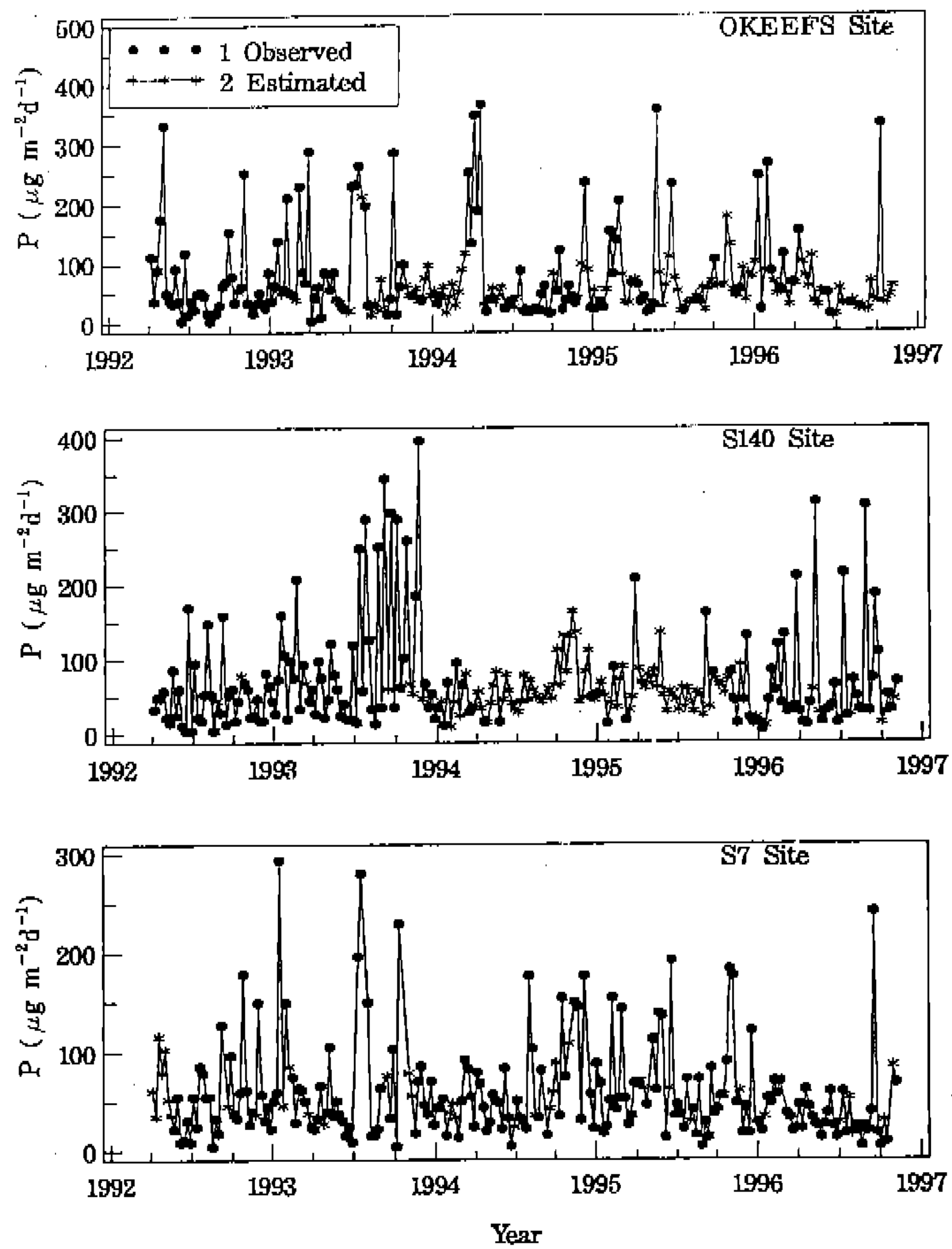
Month	Mean	Standard Deviation	N	Minimum	Maximum
January	61.4	51.0	184	4.1	294.1
February	86.2	69.1	160	5.4	363.0
March	87.1	70.6	176	12.2	354.9
April	73.8	61.4	193	4.1	336.6
May	82.2	74.8	202	3.4	361.0
June	58.6	60.5	190	2.0	352.9
July	72.8	74.4	193	4.1	416.3
August	65.8	69.8	207	4.1	354.9
September	87.2	80.5	209	4.1	346.8
October	110.8	100.7	228	4.1	449.7
November	94.3	81.2	191	4.1	363.0
December	84.1	80.2	165	4.1	365.0
Mean	80.3± 14.8+	72.9± 12.6+	192	2.0*	449.7*

List of Figures



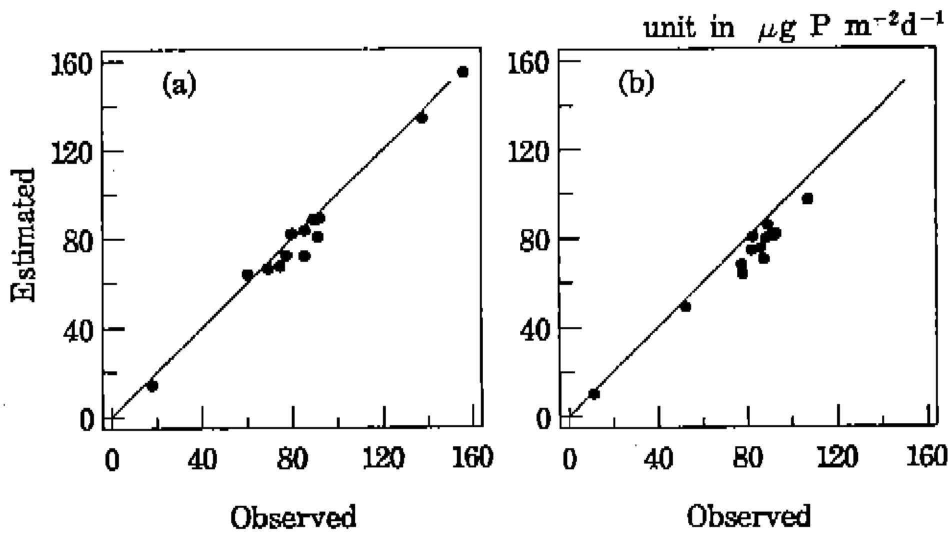
Area Map
S65A
Atlantic
Ocean
OKEEFS
G36
S127
S131
Lake
Okeechobee
S308
S310
BG1
BG2
GULF OF MEXICO
ENR
ENR101
ENR203
ENR401
ENR301
WCA 1
L6
S7
WCA 2
S140
WCA 3
L67A
ENPRC
0 10 20 30km

Fig. 1



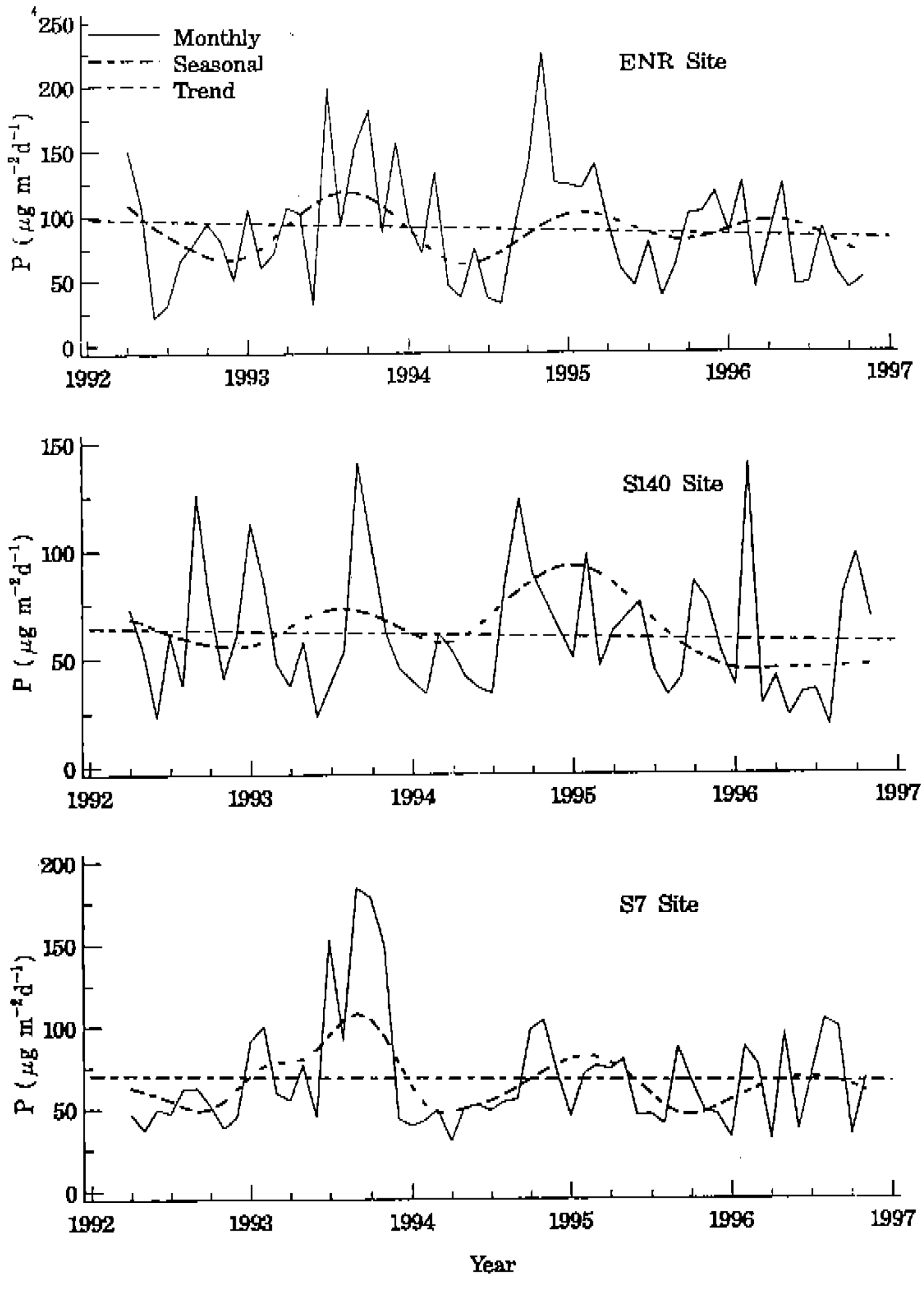
~dry/zplt2/1ts-sum

Fig. 2



~dry/zplt2/4com-sum

Fig. 3



dry/zplt2/7tr-sum

Fig. 4